PATENTTI- JA REKISTERIHALLITUS
NATIONAL BOARD OF PATENTS AND REGISTRATION

Helsinki 23.1.2009

E T U O I K E U S T O D I S T U S
P R I O R I T Y D O C U M E N T

Hakija Applicant	**Tellabs Oy** **Espoo**
Patenttihakemus nro Patent application no	**20031501 (pat.115100)**
Tekemispäivä Filing date	**14/10/2003**
Kansainvälinen luokka International class	**H04L 12/56**
Keksinnön nimitys Title of invention	**"Menetelmä ja laitteisto ruuhkanhallinnan sekä siirtoyhteyskapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä"**

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä Patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista, tiivistelmästä ja piirustuksista.

This is to certify that the annexed documents are true copies of the description, claims, abstract and drawings, originally filed with the Finnish Patent Office.

Anja Kallamaa
Apulaistarkastaja

Maksu 50 €
Fee 50 EUR

Maksu perustuu kauppa- ja teollisuusministeriön antamaan asetukseen 1142/2004 Patentti- ja rekisterihallituksen maksullisista suoritteista muutoksineen.

The fee is based on the Decree with amendments of the Ministry of Trade and Industry No. 1142/2004 concerning the chargeable services of the National Board of Patents and Registration of Finland.

Osoite: Arkadiankatu 6 A
P.O.Box 1160
FI-00101 Helsinki, FINLAND
Puhelin: 09 6939 500
Telephone: + 358 9 6939 500
Telefax: 09 6939 5328
Telefax: + 358 9 6939 5328

Menetelmä ja laitteisto ruuhkanhallinnan sekä siirtoyhteyskapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä

Keksinnön kohteena on patenttivaatimuksen 1 mukainen menetelmä ruuhkanhallinnan sekä siirtoyhteyskapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä.

Keksinnön kohteena on myös patenttivaatimuksen 8 mukainen laitteisto ruuhkanhallinnan sekä siirtoyhteyskapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä.

Tässä asiakirjassa käytetään niin tunnetun tekniikan kuin keksinnönkin kuvauksessa seuraavia lyhenteitä:

BE	Palvelunlaatuluokka sovelluksille, jotka pystyvät hyödyntämään hetkellisesti vapaana olevaa tiedonsiirtoverkon kapasiteettia mutta joille ei varata tiedonsiirtoverkon kapasiteettia (Best Effort),
CoS	Palvelunlaatuluokka (Class of Service),
DSCP	Paketin kantama tieto siitä, mihin palvelunlaatuluokkaan kyseinen paketti kuuluu (Differentiated Services Code Point),
FIFO	Aikaisemmin sisään, aikaisemmin ulos -jonokuri (First In First Out -discipline),
aG+E	Palvelunlaatuluokka sovelluksille, jotka pystyvät hyödyntämään hetkellisesti vapaana olevaa tiedonsiirtoverkon kapasiteettia ja joille varataan tietty tiedonsiirtokapasiteetti (Guaranteed rate and Best Effort),
bG+E	Samanlainen palvelunlaatuluokka kuin aG+E, mutta laatuluokassa bG + E voidaan haluttaessa käyttää erisuuruista ylitilaussuhdetta kuin laatuluokassa aG + E,
p	Palvelunlaatuluokan sisäistä aliryhmää (esim. drop precedence) ilmaiseva muuttuja,
OBR	Ylitilaussuhde (Overbooking ratio),
QoS	Palvelunlaatu (Quality of Service),
q	Palvelunlaatuluokkaa ilmaiseva muuttuja,

SFQ Start-time Fair Queuing, eräs painokerroinperusteinen vuorotusmenetelmä [1],

SLA Palvelunlaatusopimus (Service Level Agreement),

wfq Painokerroinperusteinen vuorotusmenetelmä, käytetään yleisnimenä (weighted fair queuing),

WFQ Weighted Fair Queuing, eräs painokerroinperusteinen vuorotusmenetelmä [1],

WRED Painotusperusteinen ruuhkanrajoitusmenetelmä [3, 4] (Weighted Random Early Detection).

Pakettikytkentäisessä tietoliikennejärjestelmässä on usein edullista, että siirrettävät paketit luokitellaan kuuluviksi eri palvelunlaatiluokkiin (CoS) sen mukaan, millaisia tarpeita tietoliikennepalvelua käyttävillä sovelluksilla on, ja toisaalta sen mukaan, millaisia sopimuksia palvelun laadusta (SLA) tietoliikennepalveluntarjoaja on tehnyt asiakkaidensa (loppukäyttäjien) kanssa. Esimerkiksi tavallisen puhelinsovelluksen kohdalla on olennaista, että sovelluksen tarvitsema tiedonsiirtonopeus on käytettävissä tarvittavan ajan ja siirtoviive on riittävän pieni sekä siirtoviiveen vaihtelu riittävän vähäistä. Puhelinsovelluksessa ei ole hyötyä siitä, että sovellukselle tarjottavaa tiedonsiirtonopeutta voitaisiin hetkellisesti kasvattaa, mikäli tiedonsiirtoverkon kuormitus on kyseisenä ajankohtana vähäistä. Sen sijaan esimerkiksi ladattaessa www-sivua on erittäin edullista, jos voidaan hyödyntää verkon hetkellisestikin vapaana olevaa siirtokapasiteettia täysimääräisesti.

Usein on edullista käyttää joillakin palvelunlaatuluokilla ylitilausta. Tarkastellaan tiettyä palvelunlaatuluokkaa edustavaa sovellusta, jolle palvelunlaatusopimuksella (SLA) tilataan tietty siirtonopeus [bit/s]. Tiedonsiirtoverkon edellytetään tarjoavan kyseiselle sovellukselle tilattu siirtonopeus esimerkiksi 99.99% todennäköisyydellä. Tämän vaatimuksen täyttämiseksi tiedonsiirtolinkeistä ja muista verkkoelementeistä varataan tiedonsiirtokapasiteettia [bit/s] kyseistä palvelunlaatuluokkaa käyttäville sovelluksille. Käytettäessä ylitilausta tietystä linkistä tai muusta verkkoelementistä, varattu tiedonsiirtokapasiteetti on pienempi kuin palvelunlaatusopimuksilla (SLA) kyseisestä verkon osasta tilattujen siirtonopeuksien summa. Ylitilaus luonnollisesti kasvattaa

palvelunlaatusopimuksen (SLA) loukkaamisen todennäköisyyttä. Käytännön tiedonsiirtoverkoissa on kuitenkin epätodennäköistä, että läheskään kaikki tiettyä palvelunlaatuluokkaa käyttävät loppukäyttäjät pyrkisivät samanaikaisesti hyödyntämään palvelunlaatusopimuksensa määrittämää siirtonopeutta. Ylitilaus on palveluntarjoajan kannalta kannattavaa niin kauan, kuin ylitilauksen avulla lisääntyneet loppukäyttäjiltä saatavat suoritteet (myyty siis enemmän siirtokapasiteettia) ovat suuremmat kuin palvelunlaatusopimusloukkausten lisääntymisestä aiheutuneet kustannukset. Ylitilaussuhde (Overbooking Ratio, OBR) ilmaisee tietylle liikenteelle tilattujen siirtonopeuksien summan suhdetta kyseiselle liikenteelle varattuun tiedonsiirtokapasiteettiin. Ylitilaussuhde voi olla verkkoelementtikohtainen.

Mikäli jossain palvelunlaatuluokassa käytetään ylitilausta, se tulee järjestää siten, että tietyssä palvelunlaatuluokassa käytetty ylitilaus ei heikennä palvelun laatua muissa palvelunlaatuluokissa. Palvelun laatu voi heikentyä esimerkiksi lisääntyneen pakettihävikin muodossa, kasvaneiden siirtoviiveiden ja viiveiden vaihtelun muodossa tai siten, että sovelluksen kyky hyödyntää kunakin ajankohtana verkon vapaana olevaa siirtokapasiteettia heikkenee. Ylitilauksen aiheuttama riski palvelun laadun heikentymisestä tulee kohdistua vain siihen palvelunlaatuluokkaan, jossa ylitilausta käytetään. Tässä asiakirjassa tällaiset ehdot täyttävää ylitilausta kutsutaan *hallituksi ylitilaukseksi*.

Tarkasteellaan seuraavassa tilannetta, jossa tietoliikennepalvelu tarjoaa seuraavanlaisia palvelunlaatuluokkia:

- aG + E (Guaranteed rate and Best Effort): sovellukselle, jolle palvelunlaatusopimuksella (SLA) tilataan tietty (vähimmäis)siirtonopeus [bit/s] ja jolle tarjottavaa hetkellistä tiedonsiirtonopeutta kasvatetaan hyödyntäen kunakin ajankohtana vapaana olevaa tiedonsiirtojärjestelmän kapasiteettia. aG+E palvelunlaatuluokkaa edustaville sovelluksille verkkoelementeistä varataan tiedonsiirtokapasiteettia [bit/s].

- bG + E: vastaava palvelunlaatuluokka kuin aG + E, mutta palvelunlaatuluokassa bG + E voidaan haluttaessa käyttää erisuuruista ylitilaussuhdetta (OBR_{bG+E}) kuin palvelunlaatuluokassa aG + E (OBR_{aG+E}).

- BE (Best Effort): sovelluksille, joille ei verkkoelementeistä varata tiedonsiirtokapasiteettia eikä toisaalta palvelunlaatusopimuksilla (SLA) tilata (vähimmäis)siirtonopeutta mutta joille hyödynnetään kunakin ajankohtana vapaana olevaa tiedonsiirtojärjestelmän kapasiteettia.

Kuvio 1 esittää yhtä tunnetun tekniikan mukaista tapaa vuorottaa yhteisen siirtolinkin kapasiteettia yllämainittuja palvelunlaatuluokkia (aG+E, bG+E tai BE) edustaville liikennevoille. Kuvion 1 esittämän järjestelmän toiminta on seuraava:

- Se, mihin palvelunlaatuluokkaan q yksittäinen paketti kuuluu, on identifioitavissa pakettiin liitetyn tiedon perusteella (esimerkiksi DSCP = Differentiated Services Code Point [2]).

- Paketit ohjataan palvelunlaatuluokkakohtaisiin FIFO jonoihin 3-5 (aG+E-, bG+E ja BE-jono).

- Jokainen aG+E tai bG+E palvelunlaatuluokkaa edustava paketti kuuluu palvelunlaatuluokan sisäiseen aliryhmään (p), jonka perusteella voidaan päätellä vähintään se, kuuluuko kyseinen paketti siihen osaan liikennettä, joka vastaa palvelunlaatusopimuksessa (SLA) tilattua vähimmäissiirtonopeutta (jatkossa tätä osuutta kutsutaan G-osuudeksi), vai kuuluuko paketti siihen osaan liikennettä, joka ylittää tilatun vähimmäissiirtonopeuden (jatkossa tätä osuutta kutsutaan E-osuudeksi). Tiettyyn aliryhmään p kuuluminen voidaan indikoida esimerkiksi DSCP:n kantaman etuoikeustiedon (drop precedence) avulla [2]. Aliryhmätietoa käytetään silloin, kun jonon ruuhkautuessa tulee päättää, mihin jonossa oleviin tai jonoon saapuviin paketteihin ruuhkanrajoitustoimenpiteet kohdistetaan. Esimerkkinä tästä on WRED menetelmä (Weighted Random Early Detection) [3, 4].

- Siirtolinkin kapasiteettia vuorotellaan aG+E- 3, bG+E- 4 ja BE-jonolle 5 painokerroinperusteisella vuorotusmenetelmällä (esimerkiksi SFQ [1]). Ruuhkatilanteessa siirtolinkin kapasiteetti jaetaan aG+E, bG+E, ja BE-

palvelunlaatuluokille vastaavien painokerrointen määräämissä suhteissa (W_{aG+E} : W_{bG+E} : W_{BE})

Kuviossa 1 esitetyssä vuorotusmenetelmässä painokertoimet W_{aG+E}, W_{bG+E} ja W_{BE} on valittava silmälläpitäen sitä, että palvelunlaatuluokkia aG+E ja bG+E edustavat liikenteet saavat niille varatut osuudet siirtolinkin kapasiteetista. Ongelmana kuviossa 1 esitettyssä järjestelmässä on, ettei voida edellä kuvatun vaatimuksen täyttämisen lisäksi määrätä, millä painokertoimilla palvelunlaatuluokkaa aG+E ja bG+E edustavien liikenteiden E-osuudet ja BE-liikenne kilpailevat siitä osasta siirtolinkin kapasiteettia, jota ei ole joko varattu jotain palvelunlaatuluokkaa edustavan liikenteen käyttöön tai joka on varattu muttei ole kyseisellä hetkellä varaukseen oikeutetun liikenteen käytössä.

Kuvio 2 esittää viitteessä [5] (tämän hakemuksen tekohetkellä salainen) kuvatun tunnetun tekniikan mukaista menetelmää, jossa vuorotuspainon arvo riippuu sekä laatuluokasta (q) että aliryhmästä (p). Tällöin voidaan erikseen määrätä, 1) mikä suhteelinen osuus siirtolinkin kapasiteetista annetaan kunkin palvelunlaatuluokan sille liikenneosuudelle, joka vastaa tilattua vähimmäissiirtonopeutta (aG+E:n ja bG+E:n G-osuudet) ja 2) millä painokertoimella kunkin palvelunlaatuluokan se liikenneosuus, joka ylittää tilatun vähimmäissiirtonopeuden (aG+E:n ja bG+E:n E-osuudet ja BE), kilpailee siitä osasta siirtolinkin kapasiteettia, joka ei ole tarkasteltavalla hetkellä jonkin palvelunlaatuluokan tilattua vähimmäissiirtonopeutta edustavan liikenneosuuden (aG+E:n ja bG+E:n G-osuudet) käytössä.

Kuvion 2 mukaisessa järjestelmässä varauksiin oikeutetuille liikenneosuuksille (aG+E:n ja bG+E:n G-osuudet) tulee antaa toisaalta riittävän suuret vuorotuspainot suhteessa varauksiin oikeuttamattomien liikenneosuuksien (BE, aG+E:n ja bG+E:n E-osuudet) vuorotuspainoihin, jotta voidaan varmistua siitä, että varauksiin oikeutetut liikenneosuudet saavat ruuhkatilanteessakin niille varatut siirtokapasiteettiosuudet käyttöönsä. Toisaalta taas kyseisten vuorotuspainojen tulisi olla riittävän pienet, jotta varauksiin oikeutetuissa liikenneosuuksissa käytettävä ylitilaus heikentäisi ainoastaan sen palvelunlaatuluokan suorituskykyä, jossa ylitilausta käytetään. Ongelmana kuvion 2 mukaisessa järjestelmässä on, että mainitut vuorotuspainoja koskevat vaatimukset (varausten varmistaminen, vapaan

siirtokapasiteetin jakaminen halutuissa suhteissa ja hallittu ylitilaus) ovat keskenään ristiriidattomia vain poikkeustapauksissa.

Ongelmana kuvioissa 1 ja 2 kuvatuissa menetelmissä on lisäksi se, että tilanteessa, jossa esim. aG+E laatuluokan jono 3 on ruuhkautunut kyseisessä laatuluokassa käytettävän ylitilauksen vuoksi, ruuhkanrajoitusmekanismi (esim. WRED [3, 4]) ei kykene rajoittamaan jonon pituutta vastaavalla tavalla kuin sellaisessa tilanteessa, jossa ruuhkautuminen johtuu E-osuutta edustavan liikenteen tarjonnasta. Tämä johtuu siitä, että ruuhkanrahoitusmekanismi päättää aliryhmätiedon (esim. drop precedence) perusteella, mihin paketteihin ruuhkanrajoitustoimet kohdistetaan jononpituuden ja/tai siitä johdetun suureen ylittäessä tietyn kynnysarvon. Mikäli aliryhmätieto ilmaisee paketin kuuluvan G-osuuteen, käytetään suurempaa kynnysarvoa, joka jonon pituuden tai sen johdannaisen on ylitettävä, ennen kuin ruuhkanrajoitustoimenpide kohdistetaan kyseiseen pakettiin, kuin tilanteessa, jossa tarkasteltava paketti kuuluu E-osuuteen. Ylitilausta käytettäessä jono voi ruuhkaantua jo pelkän G-osuuden vaikutuksesta. Jonon pituuden kasvaminen suurentaa siirtoviivettä ja vaikeuttaa mm. TCP-protokollan vuonohjaus- ja valvontamekenismien [6] toimintaa.

Tämän keksinnön tarkoituksena on poistaa edellä kuvatun tekniikan puutteellisuudet ja aikaansaada aivan uudentyyppinen menetelmä ja laitteisto siirtoyhteyskapasiteetin vuorottamiseksi pakettikytkentäisten tietoliikennevoiden kesken. Keksinnön kohteena on menetelmä, jolla voidaan toteuttaa vuorotin- ja ruuhkanhallintakoneisto siten, että saavutetaan seuraavat ominaisuudet:

1) Tiettyä palvelunlaatuluokkaa edustavalle liikenteelle voidaan varata tietty osuus siirtolinkin kapasiteetista, ja

2) voidaan määrätä, millä painokertoimilla kunkin palvelunlaatuluokan liikenteen se osuus, joka ylittää kyseiselle palvelunlaatuluokalle varatun osuuden siirtolinkin kapasiteetista, kilpailee siitä osasta siirtolinkin kapasiteettia, jota ei ole joko varattu jotain palvelunlaatuluokkaa edustavan liikenteen käyttöön tai joka on varattu muttei ole kyseisellä hetkellä varaukseen oikeutetun liikenteen käytössä, ja

3) voidaan käyttää ylitilausta siten, että ylitilauksesta johtuva palvelun laadun heikkeneminen kohdistuu vain siihen palvelunlaatuluokkaan, jossa ylitilausta käytetään (hallittu ylitilaus), ja

4) voidaan estää liikennevuonohjauksen (esim. TCP protokolla [6]) kannalta haitallinen jononpituuden kasvu myös ylitilauksesta johtuvassa ruuhkatilanteessa.

Keksintö perustuu siihen, että mitataan vuorotettavaksi tulevaa liikennevuota, jonka muodostavat tiettyä palvelunlaatuluokkaa edustavat jonoon saapuvat paketit tai osa kyseisistä paketeista, ja ohjataan vuorotin- (esim SFQ [1]) ja ruuhkanrajoitusmekanismin (esim. WRED [3, 4]) toimintaa mittaustuloksien perusteella.

Keksinnön mukaisen menetelmän käyttäminen pelkästään vuorotinmekanismin ohjaukseen ei estä perinteisen aliryhmätietoon (esim. drop precedence) perustuvaa ruuhkanrajoitusmenetelmän käyttöä. Keksinnön mukaisen menetelmän käyttäminen pelkästään ruuhkanrajoitusmekanismin ohjaukseen ei estä tunnetun tekniikan mukaisten vuorotusmenetelmien käyttöä.

Mittaustulos voi olla yksi luku, jonka arvo ilmaisee ohjauksessa hyödynnettävää tietoa, tai monta lukua (vektori), joiden arvot ilmaisevat hyödynnettäviä tietoja. Jatkossa mittaustulosta käsitellään yleisyyden vuoksi usean osatuloksen muodostamana vektorina.

Keksinnön mukaiselle menetelmälle on tunnusomaista se, mikä on esitetty patenttivaatimuksen 1 tunnusmerkkiosassa.

Keksinnön mukaiselle laitteistolle puolestaan on tunnusomaista se, mikä on esitetty patenttivaatimuksen 8 tunnusmerkkiosassa.

Keksinnöllä saavutetaan tunnetun tekniikan mukaisiin ratkaisuihin verrattuna se etu, että voidaan toteuttaa vuorotin- ja ruuhkanrajoituskoneisto siten, että ylitilauksesta johtuva palvelun laadun heikkeneminen kohdistuu vain siihen palvelunlaatuluokkaan, jossa

ylitilausta käytetään, ja lisäksi voidaan estää liikennevuonohjauksen kannalta haitallinen jonon pituuden kasvu myös ylitilauksesta johtuvassa ruuhkatilanteessa.

Keksintöä ryhdytään seuraavassa lähemmin tarkastelemaan oheisten kuvioiden mukaisten esimerkkien avulla.

Kuvio 1 esittää lohkokaaviona yhtä tunnetun tekniikan mukaista järjestelmää yhteisen siirtolinkin kapasiteetin vuorottamiseksi edellämainittuja palvelunlaatuluokkia (aG+E, bG+E, BE) edustaville liikennevoille.

Kuvio 2 esittää lohkokaaviona toista tunnetun tekniikan mukaista järjestelmää yhteisen siirtolinkin kapasiteetin vuorottamiseksi edellämainittuja palvelunlaatuluokkia edustaville liikennevoille.

Kuvio 3 esittää lohkokaaviona keksinnön mukaista järjestelmää yhteisen siirtolinkin kapasiteetin vuorottamiseksi edellämainittuja palvelunlaatuluokkia edustaville liikennevoille.

Keksinnön mukaisen menetelmän teoreettinen perusta käy ilmi seuraavasta tarkastelusta.

Painokerroinperusteisessa vuorotusmenetelmässä vuorottimen 1 sisääntulossa oleville paketeille muodostetaan järjestysindikaatio (esimerkiksi Start_tag SFQ menetelmässä [1]) siitä, milloin kyseinen paketti tulee eteenpäinsiirtovuoroon. Ensimmäisenä siirretään eteenpäin se paketti, jonka järjestysindikaatio on arvoltaan sellainen, joka ilmaisee aikaisinta eteenpäinsiirtohetkeä. Järjestysindikaation ei tarvitse olla sidoksissa reaaliaikaan, vaan riittää, että eri pakettien järjestysindikaatiot ovat mielekkäässä suhteessa toisiinsa nähden.

Järjestysindikaation muodostamisessa tietystä palvelunlaatuluokkajonosta tulevalle paketille käytetään kyseistä palvelunlaatuluokkaa vastaavaa painokerrointa. Mikäli jonolla J1 on suurempi painokerroin kuin jonolla J2, niin jonon J1 peräkkäisten pakettien järjestysindikaatioiden sarja suhteessa jonon J2 vastaavaan muodostuu sellaiseksi, että jono

J1 saa suuremman osuuden vuorottimen 1 ulostulon kapasiteetista.

Prioriteettiperusteisessa vuorotusmenetelmässä vuorottimen sisääntulossa oleville paketeille annetaan prioriteettiarvo. Pakettien prioriteettiarvot määräävät, mikä paketti seuraavaksi siirretään eteenpäin.

Keksinnön mukaisessa menetelmässä paketille annettava prioriteettiarvo tai paketin järjestysindikaation muodostamisessa käytettävä painokerroin ei riipu ainoastaan paketin edustamasta palvelunlaatuluokasta (jota tässä asiakirjassa ilmennetään muuttujalla q) vaan myös kyseisen palvelunlaatuluokan liikennevuohon tai kyseisen liikennevuon osaan kohdistetun mittauksen 3 antamista tuloksista (joita tässä asiakirjassa ilmennetään muuttujavektorilla x), kuvio 3.

Keksinnön mukaisessa menetelmässä mittaustulostieto/-tiedot voivat määrätä myös sen, käytetäänkö tietyn paketin vuorotuspäätösen tekoon painokerroin- vai prioriteettiperusteista vuorotusmenetelmää.

Ruuhkanhallinnassa käytetään hyväksi jonon pituutta tai siitä johdettua suuretta kuten esimerkiksi jonon pituuden alipäästösuodatettua arvoa. Mikäli jonon pituus ja/tai sen johdannainen ylittää tietyn kynnysarvon, kohdistetaan ruuhkanrajoitustoimenpiteitä tiettyihin jonossa oleviin tai jonoon saapuviin paketteihin. Ruuhkarajoitustoimenpiteitä voivat olla pakettien pudottaminen (tuhoaminen) tai merkitseminen (ECN menetelmä [2]). Tietyn palvelunlaatuluokan sisällä niiden pakettien valinta, joihin ruuhkanrajoitustoimenpiteitä kohdistetaan, perustuu tunnettua tekniikkaa edustavissa ruuhkanhallintamenetelmissä aliryhmätietoon (esim. drop precedence). Periaate on, että esim. aG+E luokan tapauksessa ruuhkanrajoitustoimenpiteitä kohdistetaan ensin niihin kyseistä palvelunlaatuluokkaa edustaviin paketteihin, jotka aliryhmätiedon perusteella kuuluvat E-osuuteen. Jollei jonon pituuden kasvu pysähdy pudottamalla (tai merkitsemällä) E-osuutta edustavia paketteja, aletaan pudottaa (merkitä) myös G-osuutta edustavia paketteja. WRED menetelmässä tämä on toteutettu siten, että G-osuudelle määrätty jonon pituuden tai siitä johdetun suureen kynnysarvo on suurempi kuin E-osuudelle määrätty vastaava kynnysarvo.

Jollei käytetä ylitilausta, pelkkien E-osuutta edustavien pakettien pudottamisen (tai merkitsemisen) tulisi jo estää ruuhkautuminen, koska G-osuutta edustavalle liikenteelle on varattu tarvittava siirtokapasiteetti. Mikäli ylitilausta käytetään, jonon kasvu voi jatkua vielä siinäkin tilanteessa, että ruuhkanrajoitustoimenpiteet kohdistuvat jo kaikkiin E-osuutta edustaviin paketteihin. Tämä johtuu siitä, että käytettäessä ylitilausta (määritelmän mukaisesti) on mahdollista, että siirtolinkille pyrkii suurempi G-osuutta edustava liikennemäärä kuin mitä G-osuutta edustavalle liikenteelle on varattu siirtokapasiteettia. Tällöin jonon pituus rajoittuu G-osuudelle määrätyn kynnysarvon perusteella. Viivekäyttäytymisen sekä vuonhallinnan (esim. TCP) kannalta on kuitenkin edullista, että jonon pituus pysyisi mahdollisimman pienenä. Tähän pyritään esim. WRED algoritmilla siten, että kynnysarvo, jonka jälkeen E-osuutta olevia paketteja aletaan pudottaa (tai merkitä), on matala. Toisaalta G-osuudella ei voida käyttää matalaa kynnysarvoa, jotta saavutettaisiin selkeä pudotus-/merkintähierarkia – ensin E-osuus ja vasta sitten kohdistetaan rajoitustoimia G-osuuteen. Näin ollen ylitilauksesta aiheutuneessa ruuhkatilanteessa esim. WRED algoritmin perustavoite pienenä pidettävästä jononpituudesta ei toteudu.

Keksinnön mukaisessa menetelmässä edellä esitetty jonon pituuteen liittyvä ongelma on ratkaistu käyttämällä ruuhkanhallinnassa aliryhmätiedon sijasta tai rinnalla mittaustuloksia **x**, kuvio 3.

Havainnollistetaan seuraavassa yhden keksinnön toteutusmuodon mukaisen vuorotin- ja ruuhkanhallintamenetelmän toimintaa aG+E ja bG+E luokkiin kuuluvien liikennevoiden osalta käyttäen SFQ vuorotusalgoritmia [1] ja WRED ruuhkanhallinta-algoritmia [3, 4]. Tässä keksinnön toteutusmuodossa pakettikohtainen painokerroin määräytyy mittaustuloksen perusteella seuraavasti:

Sille osuudelle aG+E palvelunlaatuluokkaa edustavaa liikennevuota, jolle tarkasteltavan paketin kohdalla mitattu siirrettyjen bittien määrä on mielivaltaisella tarkasteluvälillä T menneisyydestä nykyhetkeen pienempi kuin CIR × T + CBS, on pakettikohtainen painokerroin $W_{aG+E} = W_{Ga}$, yli menevälle osuudelle $W_{aG+E} = W_{Ea}$. Vastaavalla tavalla

palvelunlaatuluokassa bG+E painokerroin $W_{bG+E} = W_{Gb}$ tai W_{Eb}. CIR on palvelunlaatuluokan G-osuudelle varattu käytettävissä oleva siirtokaista (committed information rate [bit/s]), joka ylitilausta käytettäessä on pienempi kuin suurin mahdollinen G-osuutta edustavan liikenteen määrä [bit/s]. CBS on suurin sallittu purskekoko (committed burst size [bit]). Tässä kuvattu mittaus voidaan toteuttaa esim. Token Bucket – menetelmällä [7].

Kutsutaan jatkossa niiden aG+E (bG+E) palvelunlaatuluokkaan kuuluvien pakettien muodostamaa liikenteen osuutta, joille pätee $W_{aG+E} = W_{Ga}$ ($W_{bG+E} = W_{Gb}$), g-osuudeksi ja vastaavasti niiden pakettien muodostamaa liikenteen osuutta, joille pätee $W_{aG+E} = W_{Ea}$ ($W_{bG+E} = W_{Eb}$), e-osuudeksi.

aG+E-luokan paketin i ja bG+E-luokan paketin j järjestysindikaatiot ($S_{aG+E}(i)$ ja $S_{bG+E}(j)$) lasketaan seuraavasti:

$$S_{aG+E}(i) = \max \{v, S_{aG+E}(i-1) + L(i-1) / W_{aG+E}\}, \quad (1)$$

$$S_{bG+E}(j) = \max \{v, S_{bG+E}(j-1) + L(j-1) / W_{bG+E}\}, \quad (2)$$

missä L(i-1), L(j-1) on edellisen paketin koko (esimerkiksi bitteinä) ja v on kulloinkin eteenpäin siirrettävänä olevan paketin järjestysindikaatio (virtuaaliaika). Järjestysindikaatio lasketaan silloin, kun paketti saapuu SFQ-koneiston laatuluokkakohtaiseen sisääntuloon, eikä sitä päivitetä myöhemmin vaikka v muuttuisi. Seuraavaksi eteenpäinsiirrettäväksi valitaan se paketti (i tai j), jonka järjestysindikaatio on pienempi.

Yksinkertaisella kokeilulla tai simulaatiolla voidaan todeta seuraavaa: jos tietyllä aikavälillä siirrettävät palvelunlaatuluokan aG+E paketit kuuluvat aG+E:n g-osuuteen ja palvelunlaatuluokan bG+E siirrettävät paketit kuuluvat bG+E:n e-osuuteen, niin tällöin kyseisellä aikavälillä siirrettyjen aG+E ja bG+E palvelunlaatuluokkien pakettien kantamien tavujen (tai bittien) määrän suhde on $W_{Ga} : W_{Eb}$. Tarkastelu käy havainnollisemmaksi, jos kaikki paketit oletetaan samankokoisiksi. Tällöin voidaan puhua yksinkertaisesti paketeista sen sijaan että puhutaan biteistä tai tavuista edustaen paketteja. Valitsemalla painokertoimet W_{Ga}, W_{Ea}, W_{Gb} ja W_{Eb} sopivasti voidaan määrätä, montako palvelunlaatuluokan aG+E g-

tai e-osuutta edustaavaa pakettia siirretään palvelunlaatuluokan bG+E g- tai e-osuutta edustavaa pakettia kohden.

Yksi tämän toteutusmuodon variaatio saadaan aikaan siten, että $W_{Ga} = W_{Gb}$, $W_{Ea} = W_{Eb}$ ja $W_{Ga} >> W_{Ea}$ ($W_{Gb} >> W_{Eb}$) esim. $W_{Ga} = 10000 \times W_{Ea}$. Tämä vastaa itse asiassa sitä, että g-osuuksiin kuuluvat paketit vuorotetaan käytännöllisesti katson prioriteettiperiaatteella siten, että palvelunlaatuluokkien aG+E ja bG+E g-osuuksilla on keskenään tasapuolisesti vuorotteleva prioriteetti. Tämä on mahdollista, koska g-osuudet ovat rajoitettuja siten, että niiden tarvitsema siirtokaista on käytettävissä.

Tässä esitettävässä keksinnön toteutusmuodossa palvelunlaatuluokan aG+E tai bG+E sisällä niiden pakettien valinta, joihin ruuhkanrajoitustoimenpiteitä kohdistetaan, perustuu aliryhmätiedon sijasta siihen, kuuluko tarkasteltava paketti g- vai e-osuuteen. Periaate on, että ruuhkanrajoitustoimenpiteitä kohdistetaan ensin e-osuutta edustaviin paketteihin. Jollei jonon pituuden kasvu pysähdy pudottamalla (tai merkitsemällä) e-osuutta edustavia paketteja, aletaan pudottaa (merkitä) myös g-osuutta edustavia paketteja. WRED menetelmässä tämä on toteutettu siten, että g-osuudelle määrätty kynnysarvo, joka jonon pituuden tai sen johdannaisen on ylitettävä, ennen kuin aletaan pudottaa (merkitä) g-osuuteen kuuluvia paketteja, on suurempi kuin e-osuudelle määrätty vastaava kynnysarvo.

Koska palvelunlaatuluokkien aG+E ja bG+E g-osuudet ovat rajoitettuja siten, että niiden tarvitsema siirtokaista on käytettävissä, pelkkien e-osuutta edustavien pakettien pudottaminen (tai merkitsemisen) jo estää ruuhkautumisen. Näin ollen jonon pituuden ruuhkatilanteessa määrää e-osuudelle asetettu kynnysarvo, joka voidaan valita pieneksi.

Yksi tämän toteutusmuodon edullinen variaatio saadaan aikaan siten, että mittaustoiminto kohdistetaan vain G-osuuteen ja ne paketit, jotka eivät kuulu G-osuuteen, käsitellään e-osuudessa. Tällöin voidaan varmistua siitä, että mahdollisimman suuri osuus niistä paketeista, jotka kuuluvat siihen osaan liikennettä, joka vastaa palvelunlaatusopimuksissa luvattua siirtonopeutta (G-osuus), tulevat käsitellyksi g-osuudessa. Mittauksen kohdistaminen ainoastaan G-osuuteen voidaan toteuttaa aliryhmätiedon (p, esim. drop precedence) perusteella.

Viitteet:

[1] Pawan Goyal, Harric M. Vin, Haichen Cheng. *Start-time Fair Queuing: A scheduling Algorithm for Integrated Services Packet Switching Networks*. Technical Report TR-96-02, Department of Computer Sciences, University of Texas Austin.

[2] Bruce Davie, Yakov Rekhter. *MPLS Technology and Applications*. Acaqdemic Press 2000 CA U.S.A. (www.academicpress.com)

[3] Sally Floyd, Van Jacobson. *Random Early Detection Gateways for Congestion Avoidance*. Lawrence Berkeley Laboratory 1993, University of California.

[4] Internet osoiteesta: http://www.juniper.net/techcenter/techpapers/200021-01.html löytyvä kuvaus WRED algoritmista.

[5] Janne Väänänen. *Menetelmä ja Laitteisto Siirtoyhteyskapasiteetin Vuorottamiseksi Pakettikytkentäisten Tietoliikennevoiden Kesken*, Suomalainen patenttihakemus n:o 20021921, Helsinki Finland 2002.

[6] Douglas E. Comer. *Internetworking with TCP/IP, Third Edition*. Prentice Hall International Editions, U.S.A. 1995.

[7] P. F. Chimento. *Standard Token Bucket Terminology*. http://qbone.internet2.edu/bb/Traffic.pdf 2000.

Patenttivaatimukset:

1. Menetelmä ruuhkanhallinnan sekä siirtolinkkikapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä, jossa menetelmässä

- digitaalista tietoa siirretään vakio- tai vaihtuvanmittaisina paketteina,
- paketteihin liittyy tunnistetieto, jonka perusteella paketit jaetaan vähintään kahteen eri palvelunlaatuluokkaan,
- palvelunlaatuluokkatiedon perusteella kukin paketti ohjataan yhteen rinnakkaisista FIFO jonoista (3-5), joita on yksi jono kutakin palvelunlaatuluokkaa kohden,
- samaan palvelunlaatuluokkaan kuuluvat paketit muodostavat vuon (flow), jossa pakettien siirtojärjestys säilytetään,
- järjestelmästä uloslähtevän linkin tai linkkien käytettävissä olevaa kapasiteettia vuorotetaan (1) palvelunlaatuluokkakohtaisille FIFO jonoille painokerroinperusteisella vuorotusmenetelmällä, prioriteettiperusteisella vuorotusmenetelmällä tai näiden yhdistelmällä,
- palvelunlaatuluokkakohtaisten FIFO jonojen ruuhkautumista rajoitetaan pudottamalla tai merkitsemällä (ECN, Explicit Congestion Notification [2]) jonossa olevia tai jonoon saapuvia paketteja,

tunnettu siitä, että pakettikohtainen prioriteettiarvo prioriteettiperusteisessa vuorotuksessa ja/tai painokerroin painokerroinperusteisessa vuorotuksessa määräytyy muuttujan q ja muuttujavektorin x yhteisvaikutuksesta ja että tietyn palvelunlaatuluokan sisällä niiden pakettien valinta, joihin ruuhkatilanteessa pudotus tai merkintä kohdistetaan, määräytyy muuttujavektorin x vaikutuksesta, missä muuttuja q määräytyy palvelunlaatuluokasta (CoS), jota edustavaan liikenteeseen kyseinen paketti kuuluu, ja muuttujavektori x koostuu tarkasteltavan palvelunlaatuluokan liikennevuohon tai liikennevuon osuuteen kohdistetun mittauksen (2) antamista tuloksista tai kyseisistä tuloksista johdetuista suureista.

2. Patenttivaatimuksen 1 mukainen menetelmä, tunnettu siitä, että muuttujavektori **x** ilmaisee, onko siirrettyjen bittien määrä mielivaltaisella tarkasteluvälillä T menneisyydestä nykyhetkeen pienempi kuin CIR × T + CBS, missä CIR on tarkasteltavan palvelunlaatuluokan käytettävissä oleva siirtokaista (committed information rate [bit/s]) ja CBS on suurin sallittu purskekoko (committed burst size [bit]).

3. Patenttivaatimuksen 1 mukainen menetelmä, tunnettu siitä, että ainakin yhdelle palvelunlaatuluokalle pätee, että siihen kuuluviin paketteihin liittyy tunnistetieto, jonka avulla paketit jaetaan vähintään kahteen palvelunlaatuluokan sisäiseen aliryhmään (esim. drop precedence), ja kyseistä palvelunlaatuluokkaa edustavan liikennevuon se osuus, johon mittaus (2) kohdistetaan, määritetään kyseistä palvelunlaatuluokkaa edustavasta liikennevuosta aliryhmätiedon perusteella.

4. Patenttivaatimuksen 1 mukainen menetelmä, tunnettu siitä, että painokerroinperusteisena vuorotusmenetelmänä käytetään SFQ menetelmää (Start-time Fair Queuing [1]).

5. Patenttivaatimuksen 1 mukainen menetelmä, tunnettu siitä, että painokerroinperusteisena vuorotusmenetelmänä käytetään WFQ menetelmää (Weighted Fair Queuing [1]).

6. Patenttivaatimuksen 1 mukainen menetelmä, tunnettu siitä, että muuttujavektorilla **x** ohjattavana ruuhkanrajotusmenetelmänä käytetään WRED menetelmää (Weighted Random Early Detection [3, 4]).

7. Patenttivaatimusten 1 ja 2 mukainen menetelmä, tunnettu siitä, että muuttujavektorin **x** sisältämä informaatio muodostetaan Token Bucket –menetelmällä [7].

8. Laitteisto ruuhkanhallinnan sekä siirtolinkkikapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä, jossa laitteisto käsittää

-välineet digitaalista tietoa kantavien vakio- tai vaihtuvanmittaisten pakettien vastaanottamiseksi,

-välineet paketteihin liittyvän tunnistetiedon lukemiseksi, jonka perusteella paketit voidaan jakaa vähintään kahteen eri palvelunlaatuluokkaan,

-välineet pakettien jakamiseksi vähintään kahteen eri palvelunlaatuluokkaan,

-FIFO jonon kutakin palvelunlaatuluokkaan kohden,

-välineet paketin ohjaamiseksi palvelunlaatuluokkatiedon perusteella kyseistä palvelunlaatuluokkaa vastaavaan FIFO jonoon (3-5),

-vuorottimen (1) järjestelmästä uloslähtevän linkin tai linkkien käytettävissä olevan kapasiteettin vuorottamiseksi palvelunlaatuluokkakohtaisille FIFO jonoille painokerroinperusteisella vuorotusmenetelmällä, prioriteettiperusteisella vuorotusmenetelmällä tai näiden yhdistelmällä,

- välineet pakettien lähettämiseksi uloslähtevälle linkille tai linkeille vuorottimen määräämässä lähetysjärjestyksessä,

- välineet palvelunlaatuluokkakohtaisten FIFO jonojen (3-5) ruuhkautumisen rajoittamiseksi pudottamalla tai merkitsemällä (ECN, Explicit Congestion Notification [2]) jonossa olevia tai jonoon saapuvia paketteja,

tunnettu siitä, että laitteisto käsittää välineet, joiden avulla pakettikohtainen prioriteettiarvo prioriteettiperusteisessa vuorotuksessa ja/tai painokerroin painokerroinperusteisessa vuorotuksessa voidaan määrittää perustuen muuttujan q ja muuttujavektorin **x** yhteisvaikutukseen ja joiden välineiden avulla palvelunlaatuluokan sisällä niiden pakettien valinta, joihin ruuhkatilanteessa pudotus tai merkintä kohdistetaan, voidaan määrittää muuttujavektorin **x** vaikutuksesta, missä muuttuja q määräytyy palvelunlaatuluokasta (CoS), jota edustavaan liikenteeseen kyseinen paketti kuuluu, ja muuttujavektori **x** koostuu tarkasteltavan palvelunlaatuluokan liikennevuohon tai liikennevuon osuuteen kohdistetun mittauksen (2) antamista tuloksista tai kyseisistä tuloksista johdetuista suureista.

9. Patenttivaatimuksen 8 mukainen laitteisto on *tunnettu siitä*, että laitteisto käsittää välineet, joiden avulla voidaan muodostaa muuttujavektori **x**, joka ilmaisee,

onko siirrettyjen bittien määrä mielivaltaisella tarkasteluvälillä T menneisyydestä nykyhetkeen pienempi kuin CIR × T + CBS, missä CIR on tarkasteltavan palvelunlaatuluokan käytettävissä oleva siirtokaista (committed information rate [bit/s]) ja CBS on suurin sallittu purskekoko (committed burst size [bit]).

10. Patenttivaatimuksen 8 mukainen laitteisto, tunnettu siitä, että laitteisto käsittää välineet pakettiin liittyvän tunnistetiedon lukemiseksi, jonka perusteella voidaan selvittää palvelunlaatuluokan sisäinen aliryhmä, johon kyseinen paketti kuuluu, ja välineet, joiden avulla voidaan määrittää aliryhmätiedon perusteella kyseistä palvelunlaatuluokkaa edustavan liikennevuon se osuus, johon mittaus (2) kohdistetaan.

11. Patenttivaatimuksen 8 mukainen laitteisto, tunnettu siitä, että laitteisto käsittää välineet painokerroinperusteisen vuorotuksen suorittamiseksi SFQ menetelmällä (Start-time Fair Queuing [1]).

12. Patenttivaatimuksen 8 mukainen laitteisto, tunnettu siitä, että laitteisto käsittää välineet painokerroinperusteisen vuorotuksen suorittamiseksi WFQ menetelmällä (Weighted Fair Queuing [1]).

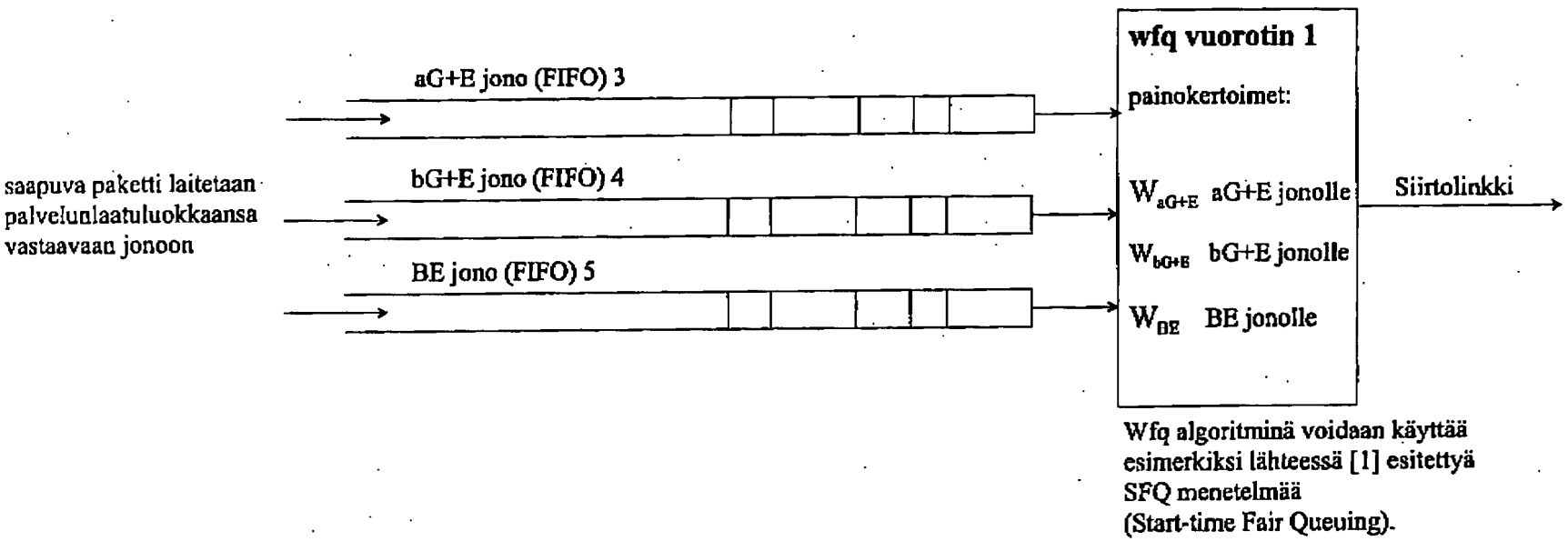
13. Patenttivaatimuksen 8 mukainen laitteisto, tunnettu siitä, että laitteisto käsittää välineet, joiden avulla muuttujavektorilla x ohjattava ruuhkanrajoittaminen voidaan suorittaa WRED menetelmällä (Weighted Random Early Detection [3, 4]).

14. Patenttivaatimusten 8 ja 9 mukainen laitteisto, tunnettu siitä, että laitteisto käsittää välineet muuttujavektorin x sisältämän informaation muodostamiseksi Token Bucket –menetelmällä [7].

Tiivistelmä:

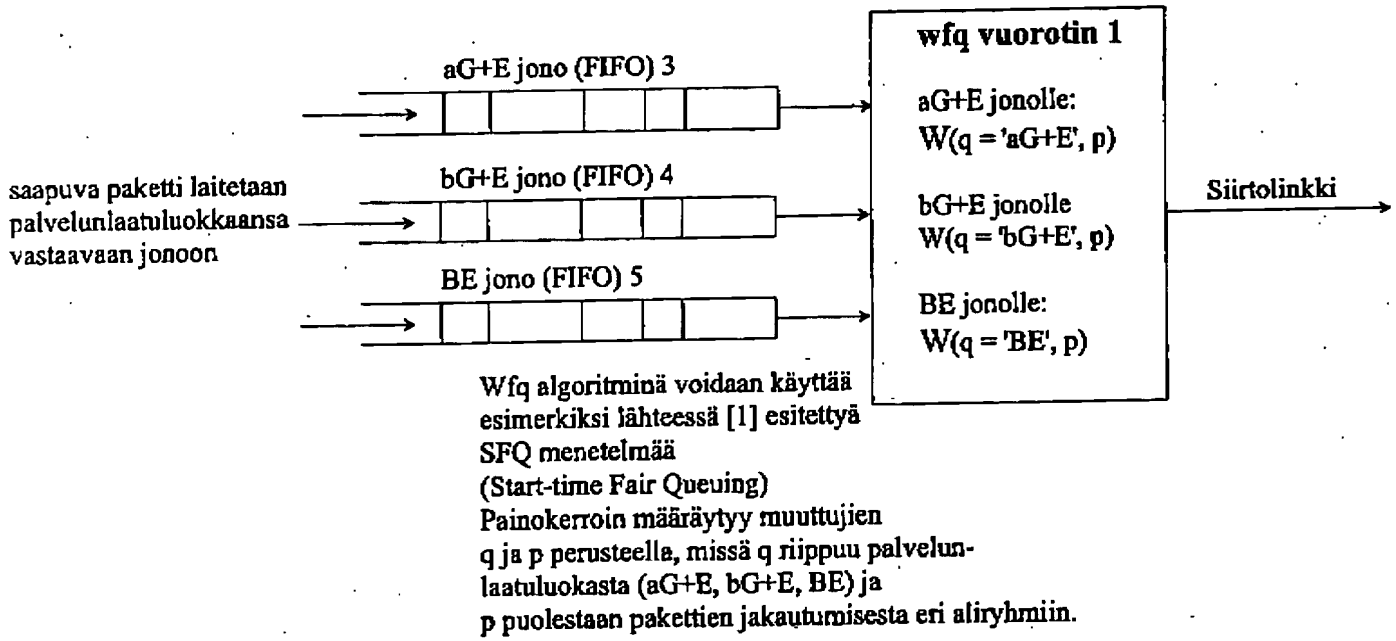
Keksinnön kohteena on menetelmä ja laitteisto ruuhkanhallinnan sekä siirtolinkkikapasiteetin vuorottamisen ohjaamiseksi pakettikytkentäisessä tietoliikenteessä siten, että 1) voidaan määrätä, mikä osuus siirtolinkin kapasiteetista varataan tiettya palvelunlaatuluokkaa edustavalle liikenteelle, ja 2) voidaan määrätä, millä painokertoimella kunkin palvelunlaatuluokan varauksen ylittävä liikenteen osuus kilpailee siitä osasta siirtolinkin kapasiteettia, jota ei ole varattu tai joka on varattu muttei ole hetkellisesti varaukseen oikeutetun liikenteen käytössä ja 3) voidaan käyttää ylitilausta siten, että ylitilauksesta johtuva palvelunlaadun heikkeneminen kohdistuu vain siihen palvelunlaatuluokkaan, jossa ylitilausta käytetään ja 4) voidaan estää liikennevuonohjauksen kannalta haitallinen viiveiden kasvu myös ylitilauksesta johtuvassa ruuhkatilanteessa. Keksintö perustuu siihen, että mitataan vuorotettavaksi tulevaa liikennevuota, jonka muodostavat tiettyä palvelunlaatuluokkaa edustavat jonoon saapuvat paketit tai osa kyseisistä paketeista, ja ohjataan vuorotin- ja ruuhkanrajoitusmekanismin toimintaa mittaustuloksien perusteella.

(Kuvio 3)



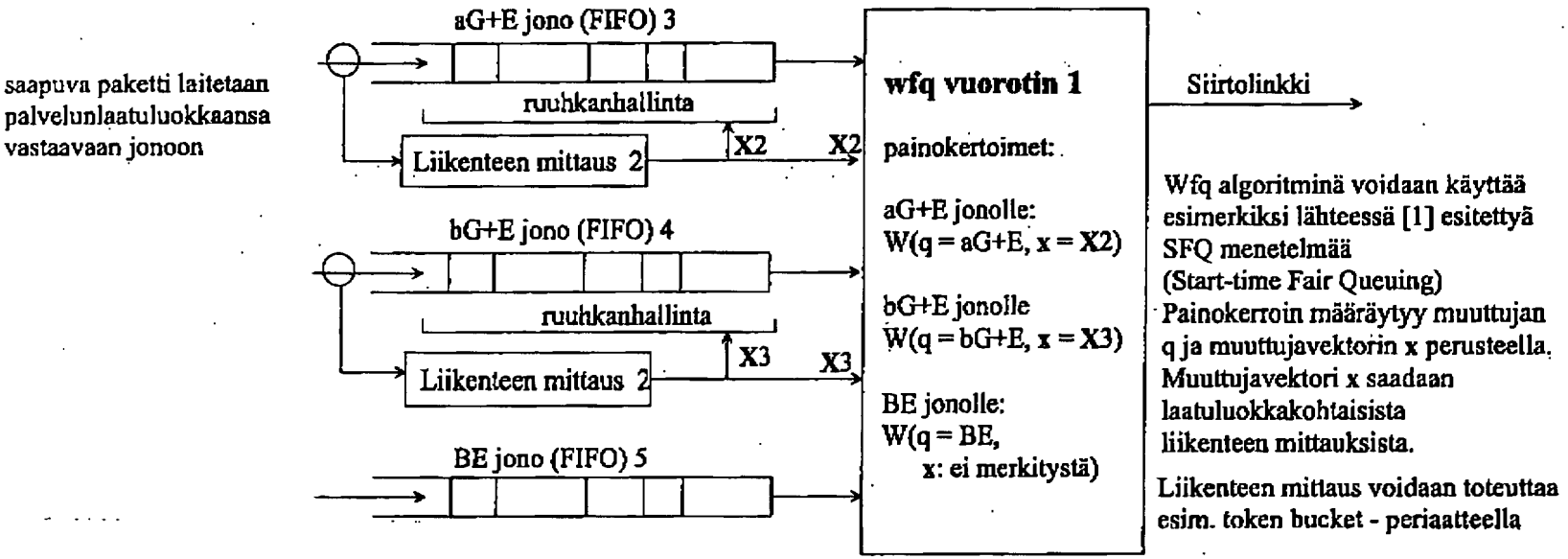
Wfq algoritminä voidaan käyttää esimerkiksi lähteessä [1] esitettyä SFQ menetelmää (Start-time Fair Queuing).

Kuvio 1



Wfq algoritminä voidaan käyttää esimerkiksi lähteessä [1] esitettyä SFQ menetelmää (Start-time Fair Queuing) Painokerroin määräytyy muuttujien q ja p perusteella, missä q riippuu palvelunlaatuluokasta (aG+E, bG+E, BE) ja p puolestaan pakettien jakautumisesta eri aliryhmiin.

Kuvio 2



Kuvio 3

CONFIRMATION

Being familiar with both the Finnish and the English languages, I herewith confirm that the attached document is a true translation of the basic Finnish Patent Application No. 20031501 filed with the Finnish Patent Office on 14 October 2003 in the name of TELLABS OY.

Helsinki on 27 January 2009

Jari Lipsanen
Itämerenkatu 3 B
FIN-00180 Helsinki
Finland

TELL11EP/P4316EP00

Method and Equipment for Controlling the Congestion Management and Scheduling of Transmission-Link Capacity in Packet-Switched Telecommunications

The present invention relates to a method, according to Claim 1, for controlling the congestion management and scheduling of transmission-link capacity in packet-switched telecommunications.

The invention also relates to equipment, according to Claim 8, for controlling the congestion management and scheduling of transmission-link capacity in packet-switched telecommunications.

In this publication, the following abbreviations are used in the descriptions of both the prior art and the invention:

BE	Service level class for applications, which are able to exploit the momentarily available capacity of a data transmission network, but for which the capacity of the data transmission network is not reserved (Best Effort),
CoS	Service level class (Class of Service),
DSCP	Data carried by a packet, stating the service level class to which the packet in question belongs (Differentiated Services Code Point),
FIFO	First In First Our discipline,
aG+E	A service level class for applications, which are able to exploit the momentarily available capacity of a data transmission network, and for which a specific data transmission capacity is reserved (Guaranteed rate and Best Effort),
bG+E	A service level class that is similar to aG+E, but, in service level class bG+E, an overbooking ratio of a different magnitude to that in service level class aG+E can be used, if desired,
[P]{p}	A variable expressing an internal sub-group (e.g., drop preference) of a service level class,
OBR	Overbooking Ratio,

QoS	Quality of service,
q	Variable expressing the service level class,
SFQ	Start-time Fair Queuing, one scheduling method [1] based on a weighting coefficient,
SLA	Service Level Agreement,
wfq	a general term (weighted fair queuing) applied to a scheduling method based on a weighting coefficient,
WFQ	Weighted Fair Queuing, one scheduling method [1] based on a weighting coefficient,
WRED	Weighted Random Early Detection, a congestion limitation method [3, 4] based on a weighting coefficient.

In a packet-switched telecommunications system, it is often preferable for the packets being transmitted to be classified as belonging to different service level classes (CoS), according to the requirements of the applications used by the telecommunications service, and, on the other hand, according to the kind of agreements on the service quality (SLA) the telecommunications service provider has made with its customers (end users). For example, in the case of normal telephone applications, it is essential for the data transmission speed required by the application to be available for the time required, for the transmission delay to be sufficiently small, and the variation in the transmission delay to be sufficient low. In telephone applications, there is no advantage in being able to momentarily increase the data transmission speed provided for an application, if the loading on the telecommunications network is small at the time in question. On the other hand, when downloading web-pages, it is extremely advantageous to be able to exploit the full, even momentarily available, transmission capacity of the network.

It is often advantageous to use overbooking for some service level classes. An application representing a specific service level class, for which a specific transmission speed [bit/s] is ordered by the service level agreement (SLA), will be examined. The telecommunications network is required to provide the transmission speed ordered for the application in question, with a probability of 99.99 %. In order to meet this demand, the data transmission capacity [bit/s] is reserved in the data transmission links and other network elements for applications using the service level class in question. When using

overbooking, the data transmission capacity reserved in a specific link or other network element is lower than the total sum of the transmission speeds ordered in the service level agreements (SLA) in the case of the relevant part of the network. Overbooking naturally increases the probability of breaching the service level agreement (SLA). However, in practice it is improbable that even nearly all of the end users using the specific service level class will attempt to simultaneously utilize the transmission speed defined in their service level agreement. From the point of view of the service provider, overbooking is profitable, as long as the payments from end users received with the aid of overbooking (thus selling more transmission capacity) are greater than the costs incurred by the increase in breaches of the service level agreements. The overbooking ratio (OBR) expresses the ratio of the total sum of the transmission speeds ordered for specific traffic to the data transmission capacity reserved for the traffic in question. The overbooking ratio can be network element specific.

If overbooking is used in some service level class, it should be arranged so that the overbooking used in the specific service level class does not reduce the quality of service in other service level classes. Service quality reduction can appear, for example, in the form of increased packet loss, of increased transmission delays and delay variations, or in a reduced ability of the application to utilize the available transmission capacity of the network at any time. The risk of service quality reduction caused by overbooking should affect only the service level class, in which the overbooking is used. In this publication, overbooking meeting such conditions is termed *controlled overbooking*.

The following examines a situation, in which the telecommunications services provides the following types of service level class:

-aG + E (Guaranteed rate and Best Effort) for an application, for which a service level agreement (SLA) is used to order a specific (minimum) transmission speed [bit/s] and for which the momentary data transmission speed provided is increased exploiting the data transmission system capacity available at each time. A data transmission capacity [bit/s] is reserved in the network elements for applications representing the aG + E service level class.

-bG + E: a service level class corresponding to aG + E, but in the service class bG + E it is possible, if desired, to use an overbooking ratio (OBR_{bG+E}) of a different magnitude to that in the service level class aG + E (OBR_{aG+E}).

-BE (Best Effort): for applications, for which a data transmission capacity is neither reserved in the network elements, nor, on the other hand, is a (minimum) transmission speed ordered using a service level agreement (SLA), but for which the telecommunications system's capacity available at any time is exploited.

Figure 1 shows one way according to the prior art of scheduling the capacity of a common transmission link for traffic flows representing the aforementioned service level classes (aG+E, bG+E, or BE). The operation of the system shown in Figure 1 is as follows:

-The service level class q, to which an individual packet belongs, can be identified on the basis of information attached to the packet (for example, DSCP = Differentiated Services Code Point [2]).

-Packets are routed into service-level-specific FIFO queues 3 - 5 (aG+E, bG+E, and BE queues).

-Each packet representing an aG+E or bG+E service level class belongs to an internal sub-group (p) in the service level class, on the basis of which it is possible to decide at least whether the packet in question belongs to the portion of the traffic that corresponds to the minimum transmission speed order in the service level agreement (SLA), (this will be subsequently referred to as the G portion), or whether the packet belongs to the portion of the traffic that exceeds the ordered minimum transmission speed (this will be subsequently referred to as the E portion). Membership of a specific sub-group p can be indicated, for example, with the aid of drop precedence information carried by the DSCP [2]. The sub-group information is used when congestion requires a decision to be made as to the packets in the queue or arriving in the queue on which congestion limitation measures should be imposed. An example of this is the WRED method

(Weighted Random Early Detection) [3, 4].

-The capacity of the transmission link is scheduled between the aG+E, bG+E, and BE queues 5, using a weighting coefficient based scheduling method (for example SFQ [1]). In a congestion situation, the capacity of the transmission link is divided between the aG+E, bG+E, and BE service level classes, in a ratio determined by the corresponding weighting coefficients ($W_{aG+E} : W_{bG+E} : W_{BE}$).

In the scheduling method shown in Figure 1, the weighting coefficients W_{aG+E}, W_{bG+E}, and W_{BE} are chosen while bearing in mind that the traffic representing the service level classes aG-E, and bG+E must receive the portions of the capacity of the transmission link reserved for them. A problem in the system shown in Figure 1 is that, in addition to meeting the aforementioned requirement, it is not possible to define the weighting coefficient by which the E portions of the traffic representing the service level classes aG+E and b+E and the BE traffic will compete for the portion of the capacity of the transmission link that is not either reserved for the use of traffic representing some service level class, or is reserved, but is not being used at the moment in question by traffic entitled to the reservation.

Figure 2 shows a method according to the prior art disclosed in reference [5] (confidential at the time of writing the present application), in which the value of the scheduling weight depends on both the quality class (q) and the sub-group (p). It is then possible to separately define 1) what relative portion of the capacity of the transmission link will be given to the portion of traffic of each service level class that corresponds to the ordered minimum transmission speed (the G portions of aG+E and bG+E) and 2) with what weighting coefficient the traffic portion, which exceeds the ordered minimum transmission speed (the E portions of aG+E and bG+E and BE), will compete for that portion of the capacity of the transmission link, which is not at the moment under examination being used by a traffic portion (the G portions of aG+E and bG+E) representing the minimum transmission speed ordered for some service level class.

In the system according to Figure 2, the traffic portions (the G portions of aG+E and bG+E) entitled to the reservations should be given, on the one hand, sufficiently large

scheduling weights relative to the scheduling weights of the traffic portions (BE and the E portions of aG+E and bG+E) unentitled to the reservations, so that it is possible to ensure that the traffic portions entitled to the reservations receive the use of the transmission capacity portions reserved for them, even in a congestion situation. On the other hand, however, the scheduling weights in questions should be small enough so that the overbooking to be used in traffic portions entitled to the reservations will reduce the performance of only the service level class in which overbooking is used. A problem in the system according to Figure 2 is that it is only in exceptional cases that the said requirements affecting the scheduling weights (ensuring reservations, division of the available transmission capacity into the desired ratios, and controlled overbooking) will not be mutually contradictory.

An additional problem in the methods shown in Figures 1 and 2 is that, in a situation in which, for example, the queue 3 of the aG+E quality class has become congested in the quality class in question due to overbooking being used, the congestion limitation mechanism (e.g., WRED [3, 4]) will not be able to limit the length of the queue in a manner corresponding to that in a situation, in which the congestion is due to traffic representing the E portion being offered. This is because the congestion limitation mechanism uses sub-group information (e.g., drop precedence) to decide which packets to apply the congestion limitation measures to, when the queue length and/or a variable derived from it exceeds a specific threshold value. If the sub-group information states that the packet belongs to the G portion, a higher threshold value is used, which the queue length or its derivative must exceed before a congestion limitation measure is applied to the packet in question, than in a situation in which the packet being examined belongs to the E portion. When using overbooking, the queue can become already congested due to the effect of only the G portion. Any increase in the length of the queue will increase the transmission delay and hampers the operation of, for instance, TCP protocol flow control and monitoring mechanisms [6].

The present invention is intended to eliminate the defects of the state of the art described above and for this purpose create an entirely new type of method and equipment for scheduling transmission link capacity between packet-switched traffic flows. The object of the invention is a method, by means of which a scheduler and congestion

management mechanism can be implemented, in such a way that the following properties are achieved:

1) A specific portion of the capacity of the transmission link can be reserved for traffic representing a specific service level class, and

2) it is possible to define the weighting coefficient by which each portion of the traffic of the service level class, which exceeds the capacity of the portion of the transmission link reserved for the service level class in question, will compete for the portion of the capacity of the transmission link, which is either not reserved for the use of traffic representing some service level class, or which is reserved but is not being used at the moment in question by traffic entitled to the reservation, and

3) it is possible to use overbooking in such a way that the reduction in the quality of the service caused by overbooking only affects the service level class, in which overbooking is used (controlled overbooking), and

4) an increase in the queue length that is detrimental in terms of the traffic-flow control (e.g., using the TCP protocol [6]) can be prevented even in a congestion situation arising from overbooking.

The invention is based on measuring the traffic flow coming to be scheduled, in which the traffic flow mentioned is formed of packets arriving in a queue representing a specific service level class, or some of the packets in question, and the operation of the scheduler (e.g., SFQ [1]) and the congestion-limitation mechanism (e.g., WRED [3, 4]) is controlled on the basis of the measurement.

The use of the method according to the invention purely to control the scheduler mechanism does not prevent the use of a traditional congestion limitation method based on sub-group information (e.g., drop precedence). Using the method according to the invention purely to control a congestion limitation mechanism does not prevent the use of scheduling methods according to the prior art.

The measurement result can be a single number, the value of which expresses information to be utilized in control, or many number (vector), the values of which express information to be utilized. In the following, the measurement result will be treated as a vector formed of several sub-results, as it is the most general approach.

The method according to the invention is characterized by what is stated in the characterizing portion of Claim 1.

The equipment according to the invention is, in turn, characterized by what is stated in the characterizing portion of Claim 8.

The use of the invention achieves the advantage over solutions according to the prior art that it is possible to implement the scheduler and congestion-limitation mechanism in such a way that the reduction in quality arising from overbooking only affects the service level class in which overbooking is used and, in addition, can prevent an increase in the length of the queue that is detrimental to traffic-flow control, even in a congestion situation arising from overbooking.

In the following, the invention is examined in greater detail with the aid of examples according to the accompanying figures.

Figure 1 shows a block diagram of one system according to the prior art, for scheduling the capacity of a common transmission link for traffic flows representing the aforementioned service level classes (aG+E, bG+E, BE).

Figure 2 shows a block diagram of a second system according to the prior art, for scheduling the capacity of a common transmission link for traffic flows representing the aforementioned service level classes.

Figure 3 shows a block diagram of a system according to the invention, for scheduling the capacity of a common transmission link for traffic flows representing the aforementioned service level classes.

The theoretical basis of the method according to the invention will become apparent from the following examination.

In the weighting-coefficient-based scheduling system, a sequence indication (for example, Start_tag SFQ in method [1]) is arranged for the packet in the input to the scheduler 1, to state when the packet in question will be in turn for forwarding. The first packet to be forwarded is that with a sequence indication value stating the earliest forwarding moment. The sequence indication need not be bound to real time, it is sufficient if the sequence indications of the packets are in a sensible relation to each other.

When forming the sequence indication, a weighting coefficient corresponding to the service level class in question is used for packets coming from a specific service level queue. If queue J1 has a greater weighting coefficient than queue J2, then the series of sequence indications of the consecutive packets of queue J1, relative to those of the corresponding ones of queue J2 is formed to be such that the queue J1 receives a larger portion of the capacity of the output of the scheduler 1.

In the priority-based scheduling system, the packets in the input of the scheduler are given a priority value. The priority values of the packets determine which packet is the next to be forwarded.

In the method according to the invention, the priority value given to the packet, or the weighting coefficient used in forming the sequence indication does not depend only on the service level class represented by the packet (which is this publication is referred to as the variable q), but also on the result (which in this publication is referred to as the variable vector x) provided by the measurement 3 made from the traffic flow of the service level class in question on from the portion of the traffic flow in question, Figure 3.

In the method according to the invention, the measurement datum/data can also determine whether the weighting coefficient or priority-based scheduling method is used

to make the scheduling decision for a specific packet.

In congestion management, the length of the queue or a variable derived from it, such as a low-pass filtered value, is utilized. If the length of the queue and/or its derivative exceeds a specific threshold value, congestion limitation measures are applied to specific packets in the queue or arriving in it. The congestion limitation measures can be packet dropping (discarding) or marking (ECN method [2]). The selection of packets within a specific service level class, to which the congestion limitation measures are applied, is based on sub-group information (e.g., drop precedence) in a congestion management method representing the prior art. The principle is that, for example, in the case of class aG+E, the congestion limitation measures are applied to the packets representing the service level class in question, which, on the basis of the sub-group information, belong to the E portion. If the increase in the length of the queue does not stop by dropping (or marking) packets representing the E portion, packets representing the G portion are also begun to be dropped (marked). In the WRED method, this is implemented in such a way that the threshold value of the queue length, or the variable derived from it defined for the G portion is greater than the corresponding threshold value defined for the E portion. Unless overbooking is being used, the dropping (or marking) of purely packets representing the E portion should already prevent congestion, as the necessary transmission capacity has been reserved for the traffic representing the G portion. If overbooking is used, the queue can continue to increase even in a situation in which congestion limitation measures are already being applied to all packets representing the E portion. This is due to the fact that, when using overbooking (as defined), it is possible for a greater amount of traffic representing the G portion, than the transmission capacity reserved for the G portion to attempt to reach the transmission link. In that case, the length of the queue is limited on the basis of the threshold value defined for the G portion. However, in terms of the delay behaviour and flow management (e.g., TCP), it is preferable for the length of the queue to remain as short as possible. This is attempted, for example, using the WRED algorithm in such a way that the threshold value, after which packets representing the E portion are begun to be dropped (or marked) is low. On the other hand, a low threshold value cannot be used for the G portion, in order to achieve a clear dropping/marking hierarchy - limitation measures are applied first of all to the E portion and only after that to the G portion. Thus, in a congestion situation

caused by overbooking, for example, the basic objective of the WRED algorithm of keeping the queue short is not met.

In the method according to the invention, the problem described above relating to the length of the queue is solved by using measurement results x, Figure 3, instead of, or along with the sub-group information in congestion management.

The following illustrates the operation of the scheduling and congestion management method according to one embodiment of the invention in the case of traffic flows belonging to the classes aG+E and bG+E, using the SFQ scheduling algorithm [1] and the WRED congestion management algorithm [3, 4]. In this embodiment of the invention, the packet-specific weighting coefficient is defined on the basis of the measurement results as follows:

For the portion of a traffic flow representing the service level class aG+E, for which, in the case of the packet being examined, the measured number of bits transmitted is, during an arbitrary examination period T from the past to the present less than CIR x T + CBS, the packet-specific weighting coefficient $W_{aG+E} = W_{Ga}$, for the excess portion $W_{aG+E} = W_{Ea}$. Correspondingly, in the service level class b+E, the weighting coefficient $W_{bG+E} = W_{Gb}$ or W_{Eb}. CIR is the available transmission band (committed information rate [bit/s]) reserved from the G portion of the service level class, which, when using overbooking is less than the largest possible amount [bit/s] of traffic representing the G portion. CBS is the largest permitted burst size [bit] (committed burst size). The measurement described here can be implemented using, for example, the Token Bucket method [7].

The portion of traffic formed of packets belonging to the aG+E (bG+E) service level class, for which $W_{bG+E} = W_{Ga}$ ($W_{bG+E} = W_{Gb}$) is valid, will subsequently be termed the g portion, and correspondingly the portion of the traffic formed of packets, for which $W_{aG+E} = W_{Ea}$ ($W_{bG+E} = W_{Eb}$), will be termed the e portion.

The sequence indications ($S_{aG+E}(i)$ and $S_{bG+E}(j)$) of an aG+E class packet i and of a bG+E class packet j are calculated as follows:

$$S_{aG+E}(i) = \max \{v, S_{aG+E}(i-1) + L(i-1)/W_{aG+E}\}, \quad (1)$$

$$S_{bG+E}(j) = \max \{v, S_{bG+E}(j-1) + L(j-1)/W_{bG+E}\}, \quad (2)$$

in which L(i -1), L(j n-1) are the size of the preceding packet (for example, in bits) and v is the sequence indication (virtual time) of the packet being forwarded at the time of inspection. The sequence indiction is calculated when the packet arrives at the quality-level-specific input of the SFQ mechanism, nor it is updated later, even if v changes. The next packet to be forwarded is selected as the packet (i or j) with the smaller sequence indication.

A simple test or simulation can be used to demonstrate the following: if the packets of service level class aG+E being transmitted during a specific period of time belong to the g portion of aG+E and the service level class bG+E packets being transmitted belong to the e portion of bG+E, then the ratio of the bytes (or bits) carried by the aG+E and bG+E service level class packets being transmitted during the period in question is W_{Ga}: W_{Eb}. The examination gives a better illustration, if all the packets are assumed to be of the same size. It is then possible to speak simple of packets, instead of speaking of packets representing bits or bytes. By selecting suitable weighting coefficients W_{Gb}, W_{Ea}, W_{Gb}, W_{Eb}, it is possible to define how many packets representing the g or e portions of the service level class aG+E are transmitted relative to the packets representing the g or e portions of the service level class bG+E.

One variation of this embodiment is created in such a way that $W_{Ga} = W_{Gb}$, $W_{Ea} = W_{Eb}$, and $W_{Ga} >> W_{Ea}$ ($W_{Gb} >> W_{Eb}$), e.g., $W_{Ga} = 10\,000 \times W_{Ea}$. In fact, this corresponds in practice to the packets belonging to the g portion being scheduling using a priority principle in such a way that the g portions of the service level classes aG+E and bG+E have a mutually equal scheduling priority. This is possible, because the g portions are limited in such a way that the transmission band they require is available.

In this embodiment of the invention which is described, the selection of the packets inside the service level class aG+E or bG+E, to which congestion limitation measures are applied, is not based on sub-group information, but instead of whether the packet

being examined belongs to the g or e portion. The principle is that congestion limitation measures are applied first of all to packets representing the e portion. If the increase in the length of the queue does not stop by dropping (or marking) packets representing the e portion, packets representing the g portion are also begun to be dropped (marked). In the WRED method, this is implemented in such a way that the threshold value defined for the g portion, which the length of the queue or a derivative of it must exceed before packets belonging to the g portion are begun to be dropped (or marked), is greater than the corresponding threshold value defined for the e portion.

Because the g portions of the service level classes aG+E and bG+E are limited in such a way that the transmission band required by them is available, the dropping (or marking) of purely the packets representing the e portions will already prevent congestion. Thus, the length of the queue in congestion situations is determined by the threshold value set for the e portion, which can be selected to be low.

One preferred variation of this embodiment is achieved in such a way that the measuring function is applied only to the G portion and the packets that do not belong to the G portion are processed in the e portion. Thus, it is possible to ensure that the greatest possible share of the packet that belong to that portion of the traffic, which corresponds to the transmission speed promised in the service level agreement (G portion), will be processed in the g portion. The application of the measurement to only the G portion can be implemented on the basis of sub-group information (p, e.g., drop precedence).

References:

[1] Pawan Goyal, Harric M. Vin, Haichen Cheng. *Start-time Fair Queuing: A Scheduling Algorithm for Integrated Services Packet Switching Networks.* Technical Report TR-96-02, Department of Computer Sciences, University of Texas Austin.

[2] Bruce Davie, Yakov Rekhter. *MPLS Technology and Applications.* Academic Press 2000 CA U.S.A. (www.academic.press.com)

[3] Sally Floyd, Van Jacobson. *Random Early Detection Gateways for Congestion*

Avoidance. Lawrence Berkeley Laboratory 1993, University of California.

[4] A description of the WRED algorithm can be found at the Internet address: http://www.juniper.net/techncenter/techpapers/200021-01.html.

[5] Janne Väänänen. *Menetelmä ja Laitteisto Siirtoyhteyskapasiteetin Vuorottamiseksi Packettikytkentäisten Tietoliikennevoiden Kesken (Method and Equipment for Sequencing Transmission Capacity Between Packet-Switched Data Traffic Flows)*, Finnish patent application No. 20021921, Helsinki Finland 2002.

[6] Douglas E. Comer. *Internetworking with TCP/IP, Third Edition*. Prentice Hall International Editions, U.S.A. 1995.

[7] P.F. Chimento. *Standard Token Bucket Terminology*. http://qbone.internet2.edu/bb/Traffic.pdf 2000.

Claims:

1. A method for controlling the congestion management and the scheduling of transmission link capacity in packet-switched telecommunications, in which method

- digital information is transmitted as constant or variable-length packets,
- identifier data is attached to the packets, on the basis of which the packets are divided into at least two different service level classes,
- on the basis of the service level class data, each packet is routed to one of the FIFO queues (3 - 5), which are one for each service level class,
- the packets belonging to the same service level class form a flow, in which the transmission order of the packets is retained,
- the available capacity of the outgoing link or links of the system is scheduled (1) for the service-level-class-specific FIFO queues using a weighting-coefficient-based scheduling method, a priority-based [sequencing] {scheduling} method, or a combination of these methods,
- congestion in the service-level-class-specific FIFO queues is limited by dropping or marking (ECN, Explicit Congestion Notification [2]) packets in the queue or arriving in the queue,

c h a r a c t e r i z e d in that the packet-specific priority value in the priority-based scheduling and/or the weighting coefficient in the weighting-coefficient-based scheduling is defined from the joint effect of a variable q and a variable vector x and that the selection of the packets within a specific service level class, to which dropping or marking will be applied in a congestion situation, are defined from the effect of the variable vector x, in which the variable q is defined from the service level class (CoS), to which the traffic represented by which the packet in question belongs, and the variable vector x is formed of the results provided by measurement (2) applied to the traffic flow representing the service level class being examined ., or of variables derived from the relevant results, in which the measurement results depend on temporal variation in the data transmission speed of the traffic representing the traffic flow being examined.

2. The method according to Claim 1, characterized in that the temporal variation in the

data transmission speed is depicted using a double-value variable, which states whether the number of bits transmitted during an arbitrary monitoring interval T from the past to the present is less than CIR x T + CBS, in which CIR is the transmission band available to the service level class being examined (committed information rate [bit/s]) and CBS is the greatest permitted burst size (committed burst size [bit/s]).

3. The method according to Claim 1, characterized in that at least one service level class is such that identifier data is attached to the packets belonging to it, with the aid of which the packets are divided into at least two internal sub-groups (.e.g., drop precedence) in the service level class, and the traffic flow being examined representing the mentioned service level class, is defined from the traffic flow representing the service level class based on the sub-group data.

4. The method according to Claim 1, characterized in that the SFQ (Start-time Fair Queuing [1]) method is used as the weighting-coefficient-based scheduling method.

5. The method according to Claim 1, characterized in that the WFQ (Weighted Fair Queuing [1]) method is used as the weighting-coefficient-based scheduling method.

6. The method according to Claim 1, characterized in that the WRED (Weighted Random Early Detection [3, 4]) method is used as the congestion limitation method controlled by the variable vector x.

7. The method according to Claims 1 and 2, characterized in that the information contained in the variable vector x is formed using the Token Bucket method [7].

8. Equipment for controlling the congestion management and scheduling of transmission link capacity in packet-switched telecommunications, in which the equipment includes

- means for receiving constant or variable-length packets carrying digital information,
- means for reading the identifier data attached to the packets, on the basis of which the packets can be divided into at least two different service level classes,
- means for dividing the packets into at least two different service level classes,

- a FIFO queue for each of the service level classes,
- means for routing a packet in the FIFO queue (3 - 5) corresponding the relevant service level class, on the basis of the service level class data,
- a scheduler (1) for scheduling the capacity available to the outgoing link or links from the system to the service-level-class-specific FIFO queues, using a weighting-coefficient-based scheduling method, a priority-based scheduling method, or a combination of these,
- means for sending packets to the outgoing link or links, in a transmission order defined by the scheduler,
- means for limiting the congestion of the service-level-class-specific FIFO queues (3 - 5), by dropping or marking (ECN, Explicit Congestion Notification [2]) packet in a queue or arriving in a queue,

characterized in that the equipment includes means, with the aid of which a packet-specific priority value can be defined in priority-based scheduling and/or a weighting coefficient can be defined in weighting-coefficient-based scheduling, on the basis of the joint effect of a variable q and a variable vector x, and with the aid of which means the selection of the packets within the service level class, to which dropping or marking is applied in a congestion situation, can be defined from the effect of the variable vector x, in which the variable q is defined from the service level class (CoS), to which the traffic represented by which the packet in question belongs, and the variable vector x is formed of the results provided by measurement (2) applied to the traffic flow representing the service level class being examined, or of variables derived from the relevant results.

9. The equipment according to Claim 8, characterized in that the equipment includes means, with the aid of which a double-value variable can be formed, which states whether the number of bits transmitted during an arbitrary monitoring interval T from the past to the present is less than CIR x T + CBS, in which CIR is the transmission band available to the service level class being examined (committed information rate [bit/s]) and CBS is the greatest permitted burst size (committed burst size [bit/s]).

10. The equipment according to Claim 8, characterized in that the equipment comprises

means for reading the identifier data is attached to the packets on the basis of which the sub-group inside the service level class can be determined, into which the packet belongs, and means for determining on the basis of the sub-group information the traffic flow being examined (2) representing the mentioned service level class.

11. The equipment according to Claim 8, characterized in that the equipment includes means for performing weighting-coefficient-based scheduling using the SFQ (Start-time Fair Queuing [1]) method.

12. The equipment according to Claim 8, characterized in that the equipment includes means for performing weighting-coefficient-based scheduling using the WFQ (Weighted Fair Queuing [1]) method.

13. The equipment according to Claim 8, characterized in that the equipment includes means, with the aid of which congestion limitation controlled using the variable vector x can be performed using the WRED (Weighted Random Early Detection [3, 4]) method.

14. The equipment according to Claims 8 and 9, characterized in that the equipment includes means for forming the information contained in the variable vector x using the Token Bucket method [7].

(57) Abstract:

The invention relates to a method and equipment for controlling the congestion management and transmission-link-capacity scheduling in packet-switched telecommunications, in such a way that 1) it is possible to define what share of the capacity of the transmission link will be reserved for traffic representing a specific service level class, and 2) it is possible to define the weighting coefficient that the portion of the traffic exceeding the reservation of each service level class will use to compete for the portion of the capacity of the transmission link that is not reserved, or that is reserved but is not being used momentarily by traffic entitled to the reservation, and 3) it is possible to use overbooking, in such a way that the reduction in service quality due to overbooking affects only the service level class in which overbooking is used, and 4) it is possible to prevent an increase in delays detrimental to traffic-flow control even in a congestion situation arising from overbooking. The invention is based on measuring the traffic flow that comes to be scheduled, in which the flow is formed of packets representing a specific service level class arriving in the queue, or some of the relevant packets, and on controlling the operation of the scheduler and congestion limitation mechanism on the basis of the measurement results.

(Figure 3)